Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020)
IEEE Xplore Part Number: CFP20J32-ART; ISBN: 978-1-7281-5518-0

Anisotropic Smoothing for Illumination Invariant Face Anti-spoofing

Raghavendra R J
Department of Computer Science Engineering
Jawaharlal Nehru National College of Engineering
Shimoga, Karnataka, INDIA
raghavendra.r.j@jnnce.ac.in

R Sanjeev Kunte
Department of Computer Science Engineering
Jawaharlal Nehru National College of Engineering
Shimoga, Karnataka, INDIA
sanjeevkunte@jnnce.ac.in

Abstract - Presently, in various biometric-based access systems face recognition has been used. But few of the systems have the facility to discriminate real and spoof faces. Moreover, plenty of researchers on face live detection primarily concentrated on intrusive approaches, which are non-user approachable in nature. An innovative non-intrusive face spoofing detection technique based on HOG and ELTP descriptors is proposed in this paper. More precisely, this approach adopts anisotropic smoothing for normalizing the illumination of a given image. The effectiveness of the proposed method has been evaluated by different types of experiments on the NUAA dataset.

Index Terms - Anisotropic smoothing, Histogram of Oriented Gradients, Extended Local Ternary Pattern.

I. INTRODUCTION

From the last three decades, there has been a lot of improvement in research on face recognition [1]. State of the art (SOTA) techniques perform well over renowned datasets and beats the human-being [2]. As a result of that, face recognition systems are used in various real-world biometric-based access applications such as face authentication in mobiles, management of identity cards and verification of security portal. Face recognition system depends on private data, but they are existed at social media (public) or can be effortlessly captured. These data can be utilized for spoofing attacks of the face recognition system. In a spoofing attack, an intruder, act as if to be a real user, attempts to gain personal information. Photos, videos, and recorded videos or 3D masks are included as common types of spoofing attacks.

By examining the flexibility of many face recognition systems against spoofing attacks [3, 4], the results represent that all the tested systems are vulnerable. Consequently, there is a vital need to find a solution to identify the spoofing attacks and safeguard the face recognition system.

Numerous face anti-spoofing approaches have been recommended recently, the detailed survey of all these approaches is presented in [5]. Based on the fact that causes of the reflection of lights from planar spoofing medium by defects in printing and artifacts, the texture-based approaches able to discriminate between real and spoof images. Many approaches use a single texture descriptor to classify the real and spoofed faces. However, Fig. 1 shows a face image which

Fig. 1: Demonstration of different image resolution and quality (low, medium, and high) of face images of NUAA face anti-spoofing dataset.

can have different image quality and resolutions. As a consequence, a single texture descriptor cannot encode them well with different types of image resolutions and produces substandard results. To deal with this type of problem Maatta et al. [6] used multiple texture descriptors instead of single texture descriptors.

In proposed work, the problem in terms of face image resolution and image quality are addressed from a different point of view. It has been proposed to represent the images in anisotropic smoothing and then extracts features from the filtered images. Anisotropic smoothing can act as a pre-processor, which increases the robustness of the face image representation against noise and illumination. To extract the features, a new descriptor ELTP+HOG (Extended Local Ternary Pattern and Histogram of Oriented Gradients) has been experimented. Many experiments are conducted on NUAA face anti-spoofing datasets and the results are compared with existing work.

The rest of the parts of the paper is structured as follows. Section II briefs existing work on face anti-spoofing methods. In section III, Anisotropic smoothing is presented and a description of HOG and ELTP descriptors is provided. In section IV, results and discussion of the proposed method are presented. Finally, the conclusion of the work carried out in section V.

II. RELATED WORK

In this section, present liveness detection approaches are discussed in brief. Two different versions of anti-spoofing approaches are discussed below.

a. Hardware-based approaches: With the advancement of electronics devices, innovative sensors are produced to obtain a variety of evidence such as LFC [7], EEG [8], Kinect [9], flashlight [10], etc. hardware-based approaches depend on the innovative devices to differentiate real and attack faces.
b. Software-based approaches: These approaches are very frequently used in face anti-spoofing systems. Software-based approaches are divided into two main streams depends on the user's cooperation. If the user does any action such as the movement of the mouth or blinking eyes for face authentication, called intrusive approaches, otherwise, called non-intrusive approaches.
   i. Intrusive approaches: These approaches are based on motion analysis, which uses such as few, eye-blinking [11, 12], movement of mouth [11], rotation of head [13], etc, in which clients need to act specifically as per instructions.
   ii. Non-intrusive approaches: User's intervention is not required, purely based on the image analysis, which includes image color analysis [14], analysis based on videos [15-17], information of texture analysis [18] and analysis of image quality [19].

An overview of some of the existing feature descriptors that are used in face spoofing detection is presented here. Ojala et al. introduced a texture descriptor called LBP (local binary pattern). LBP is not robust against illumination changes and noise in images. Tan and Triggs [20] introduced local ternary pattern (LTP), which is robust against noise. But LTP has limitations to set right threshold values for a specific application. Later, Eimad et al. introduced a local graph structure (LGS), which extracts more spatial information compared to the above descriptors. However, LGS maintains unbalanced graph structures and extracts unequal spatial information. To overcome the above deficiency, an ELTP+HOG descriptor with anisotropic smoothing is proposed in this work.

III. PROPOSED WORK

The flow of our proposed method is depicted in Fig. 2. At the pre-processing stage, anisotropic smoothing is applied to the given input image to normalize the illumination. To extract the features first, ELTP+HOG descriptors are applied to each normalized image separately and calculated the histogram of it. Then resultant histograms are concatenated to formulate the final feature vector of ELTP+HOG. Lastly, trained SVM will discriminate the given test image is real or spoof.

The details of the Anisotropic smoothing and ELTP+ HOG features are given in the following section.

A. *Anisotropic smoothing*

Anisotropic smoothing [21] is a simple automatic normalization algorithm used in image processing. It begins with evaluating the illumination that exists in the field and later, compensates by improving the local contrast of the image similarly as the human eye. This technique based on two assumptions. 1) Human vision is insensitive to illumination and sensitive to scene reflectance and 2) human vision is passive to global brightness levels and active to local changes in contrast. The above assumptions are co-related because local contrast is a reflectance function.

By using these assumptions, L(x,y) has to be found out to enhance the local contrast suitably by using eq. (1).

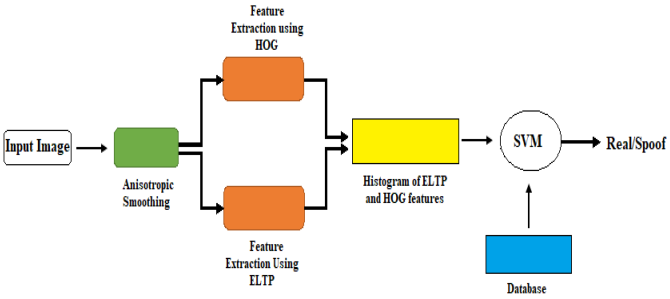

Fig. 2: Architecture of ELTP+HOG method

$$I(x,y)\frac{1}{R(x,y)} = L(x,y) \qquad (1)$$

Where, R(x,y) is the perceived sensation (reflectance of the scene), L(x,y) is the perception gain (illumination field) and I(x,y) is the input stimulus.

B. *Histogram of Oriented Gradients (HOG)*

The essential principal of the Histogram of oriented gradients is by segmenting the image into small cells (spatial regions). Each cell is represented by the combination of a local 1-D histogram of gradient directions/ edge directions over the pixels.

In the first step, calculate the magnitude of the chosen region of interest of an image by using horizontal gradient $\partial_h(x_i, y_i)$ and vertical gradient $\partial_v(x_i, y_i)$. Later, gradients are converted to polar coordinates, with an angle $\theta(x_i, y_i)$ lies between 0^o to 180^o, hence, gradients that point in opposite directions are recognized.

So,

$$f(x_i,y_i) = \sqrt{(\partial_h(x_i,y_i)^2 + \partial_v(x_i,y_i)^2)} \qquad (2)$$

$$\theta(x_i,y_i) = arctan\left(\partial_v(x_i,y_i) \Big/ \partial_h(x_i,y_i)\right) \qquad (3)$$

In the second step, the orientation histogram from the magnitude and orientations of HOG feature extraction is derived. A single window method is used in HOG feature

extraction. In a given image, the selected region of interest is divided into blocks and then blocks are segmented into cells. Later, from each cell one histogram is extracted and combined with all the extracted cells. These cells can be either in the shape of rectangular (R-HOG) or circular (C-HOG). All histogram comprises the same number of bins and represents the gradient orientations θ(xi, yi), which lies between 0^o to 180^o for unsigned gradients and 0^o to 360^o for signed gradients.

C. *Extended Local Ternary Pattern (ELTP)*

A new feature descriptor called ELTP is introduced into use for the detection of face spoofing [22]. This descriptor is invariant to the transformation of gray-levels. Each pixel's gray-level of an image is restricted depending on the tolerance zone width called ±Z around the center pixel g_c. The value is zero if it is within g_c ±Z, beyond the g_c ±Z the value is quantized to +1 else it is -1. ELTP is represented with a function containing ternary values as shown in equation (4).

$$f(g_p, g_c, z) = \begin{cases} 1 & g_p - g_c \geq Z \\ 0 & |g_p - g_c| < Z, \\ -1 & g_p - g_c \leq Z \end{cases} \quad p = 0,1 \ldots M-1 \qquad (4)$$

$$g_c \text{ is center pixel of } X, \quad Z = MAD(X), \quad X = \{g_k \mid k = 0,1,\ldots,8\}$$

Where, p denotes the number of neighbor pixels, gray values of neighbor pixels are denoted by g_p(p=0,1…M-1), X denotes a set of pixel values of 3x3 local region and MAD(X) represents median absolute of deviation of set X 's gray-values. In LTP fixed threshold is used. Because of this, LTP is not rotational invariance. But, in our work, it has been proposed to use a new strategy in ELTP, which uses dynamic threshold by auto-adaptive approach. The calculation of threshold values of ELTP is done by MAD for all 3x3 blocks.

To ease the work, the ELTP ternary pattern is segmented into above_half (ELTP_A) and below_half (ELTP_B) as presented in Fig.3. The final ELTP value is obtained by combining a histogram of above and below half. The descriptor ELTP is represented by eq. (5).

$$ELTP\_A_{q,s} = \sum_{0}^{p-1} 2^p \left(J\left(f(g_p, g_c, Z), 1\right)\right)$$

$$ELTP\_B_{q,s} = \sum_{0}^{p-1} 2^p \left(J\left(f(g_p, g_c, Z), -1\right)\right)$$

$$J(x,y) = \begin{cases} 1, & x = y \\ 0, & x \neq y \end{cases} \qquad (5)$$

D. *Classification*

The SVM classifier is used to classify the images. In all our experiments, calculation of histograms is done for both the descriptors, they are combined and the SVM classifier is trained for discrimination of given a test-image as real or attack. The following algorithm used.

Algorithm:

1. From NUAA dataset load the input-images.
2. Transform the RGB input-image of size 160 x 160 into a gray-scale image.
3. Add Anisotropic smoothing to grayscale images to normalize for illumination.
4. Segment the original image into no. of blocks of different size (40 x 40, 80 x 80).
5. Extract 128 features from each block by applying the HOG descriptor.
6. Extract 256 features from each block by applying the ELTP descriptor.
7. Histogram features are obtained from combined features of HOG and features of ELTP.
8. Classify the given test-image by using Support Vector Machine (SVM) to discriminate the real and spoof faces.

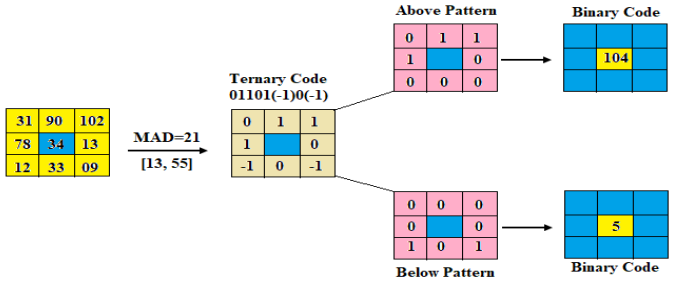

Fig. 3: ELTP Calculation Process

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

Three different types of experiments on the NUAA database [23] comprises 14 different clients of real plus attack images are investigated. Features of HOG and ELTP are extracted after anisotropic smoothing. These extracted features are used to train the SVM classifier. The trained SVM is classify given a test-image as real/attack

To experiment with the proposed method, NUAA dataset is used. The NUAA dataset comprises images of 12614 of both real and spoof faces. These images were collected in two weeks with three sessions. Totally 15 subjects were used and each subject contains at least 500 images. Each image has a resolution of 640x480 pixels. In the proposed method, 14 subjects are used and shared in training and testing.

Experiment 1: ELTP+HOG comparison in terms of ARR

In this experiment, three different sizes of images ranging from one 160 x 160 (original size), four 80 x 80(block size), and sixteen 40 x 40 (block size) are considered. The training set comprises 10 categories of both real and spoof images. Category-1 contains 10 images of each class (total of 14 subjects) were considered. Subsequently, in category-2, 20 images of each class were considered and continue till

category-10. SVM is trained with category-1 of training-set and then tested. Subsequently, category-2 to category-10 were considered and then tested. The testing-set is encompassed real plus attack images of both and a total of 3099 images. Table I lists the details of the average recognition rate (ARR) of all three block sizes for the four descriptors. 384 features of ELTP and HOG are extracted. Some experiments are carried out with LBP, LTP, and LGS. It has been witnessed from Fig. 1 and TABLE 1 that increase in no. of images for training increase in recognition rate percentage also.

Experiment 2: Evaluation of ELTP+HOG with Average Precision, Average Recall, EER and HTER

In this experiment, the original image is segmented into 40x40 pixels. Training-set comprises 10 categories of both real and attack images. SVM is trained with category-1, which contains 10 images in each class (14 subjects) and then tested. Consequently for training no. of images was enlarged in ascending order from to 20 to 100 in each class from category-2 to category-10 respectively. The testing image set contains a total of 3099 images of both bonafide and attack images. TABLE 2 lists the comparison of the developed method (ELTP+HOG) and other existing approaches on the NUAA dataset.

TABLE 1
ARR (IN %) OF SEVERAL DESCRIPTORS ON NUAA DATABASE

Sl. No.	LBP	LTP	LGS	HOG_ELTP
1	56.97	79.70	66.67	89.08
2	67.57	81.50	66.77	90.27
3	67.84	83.24	70.70	91.30
4	68.30	83.40	71.00	93.03
5	71.00	83.40	76.57	94.03
6	71.80	83.50	79.07	94.57
7	75.34	83.90	83.60	95.03
8	77.77	84.10	87.17	95.63
9	86.97	88.93	87.07	96.30
10	88.17	89.67	89.37	97.03

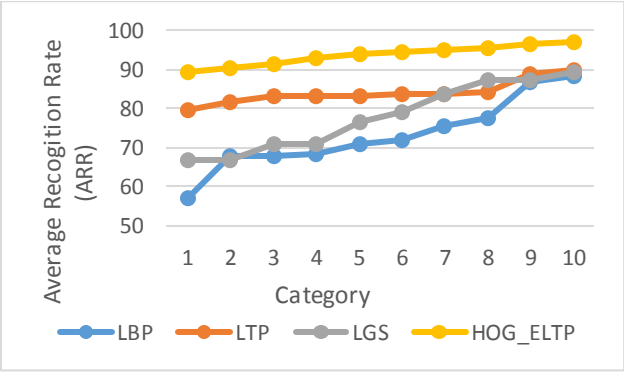

Fig. 4: Average Recognition Rate (ARR) of ELTP+HOG with other descriptors.

The results attained show that ELTP+HOG rises liveness detection of face performance over LGS, LBP, and LTP by lowermost EER (2.50%), HTER (2.59%) using an SVM classifier. ELTP+HOG outperforms over LBP by 20.44%, 27.73%, the LTP by 9.98%, 14.26%, and the LGS by 16.27%, 22.42% in terms of metrics AP (Average Precision) and (AR) Average Recall respectively as shown in Fig. 5.

TABLE 2
AP, AR, EER, AND HTER OF SEVERAL DESCRIPTORS

Descriptor	AP (%)	AR (%)	EER	HTER
LBP	74.43	63.71	10.54	12.09
LTP	84.89	77.18	9.70	11.00
LGS	78.60	69.02	9.58	10.83
ELTP+HOG	**94.87**	**91.44**	**2.50**	**2.59**

Experiment 3: Evaluation of ELTP+HOG on NUAA database

TABLE 3 lists the results of the other existing approaches on the NUAA dataset, and as a common metric HTER was used. NUAA dataset delivers a moderately no. of stable real and attack images and circumvents limited results with the HTER metric. It can be witnessed in Table 3, our method ELTP+HOG outclassed, compared to Maata J et al. [24] and X. Tan et al [25]. The HTER metric clearly shows that ELTP+HOG achieved the finest performance on the NUAA dataset. Using metric HTER, ELTP+HOG attained very good performance on the NUAA dataset.

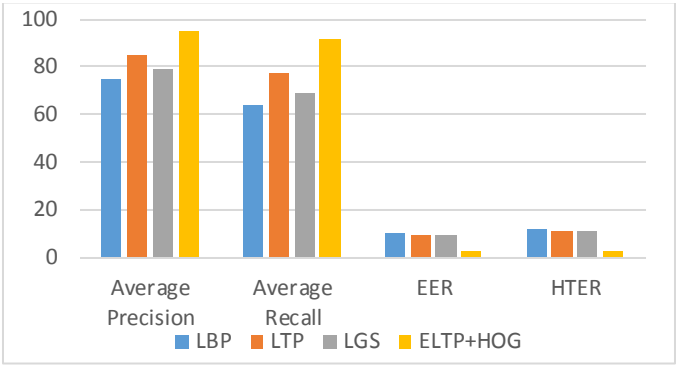

Fig.5: Performance of ELTP+HOG and other methods in terms of different metrics.

TABLE 3
RESULTS ON NUAA DATABASE

Reference	Descriptors	Classifiers	HTER (%)
Maatta J *et al.*[24]	MSLBP	SVM	19.03
Tan et al. [25]	SPARSE LOGISTIC REGRESSION MODEL	SVM	5.00
Proposed method	**ELTP+HOG**	**SVM**	**2.59**

V. CONCLUSIONS

In this work, a novel non-intrusive face spoofing detection technique based on ELTP and HOG descriptors are developed. This method used anisotropic smoothing, which acts as an

illumination invariant of a given input image. Three different experiments conducted on the NUAA dataset and acquired results are compared with various descriptors (LBP, LTP, and LGS). It has been witnessed from the experiments that ELTP+HOG performs superior to other existing methods.

REFERENCES

[1] Z. Stan, J. Anil, Encyclopedia of Biometrics, Springer, 2015.
[2] C. Lu, X. Tang, "Surpassing human-level face verification performance on LFW with Gaussian face", In Proceedings of International Conference on Artificial Intelligence, pp. 3811-3819, 2014.
[3] Y. Li, K. Xu, Q. Yan, Y. Li, R. H. Deng, " Understanding osn-based facial disclosure against face authentication systems", In Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security, ACM, pp. 413–424, 2014.
[4] L. Omar, I. Ivrissimtzis, "Evaluating the resilience of face recognition systems against malicious attacks", In Proceedings of the Computer Vision Student Workshop, pp. 1–9. 2015.
[5] J. Galbally, S. Marcel, J. Fierrez, "Biometric antispoofing methods: A survey in face recognition", In Proceedings of IEEE Access, vol. 2, pp.1530–1552, 2015.
[6] Boulkenafet Z, Komulainen J & Hadid A, "Face spoofing detection using colour texture analysis", In Proceedings of IEEE Transactions on Information Forensics & Security vol. 11, no.8, pp. 1818-1830, 2016.
[7] S. Kim, Y. Ban, S. Lee, "Face Liveliness Detection using a Light Field Camera", In Proceedings of Journal of Sensors, vol. 14, pp. 22471-22499, 2014.
[8] P. Campisi, D. L. Rocca "Brain Waves for Automatic Biometric-Based User Recognition", In Proceedings of IEEE Transactions on Information Forensics and Security, vol. 9, no. 5, pp. 782-800, 2014.
[9] Y. Wang, F. Nian, T. Li, Z. Meng, K. Wang. "Robust face antispoofing with depth information", In Proceedings of Journal of Visual Communication & Image Representation, vol. 49, pp. 332-337, 2017.
[10] P. P. K. Chan, W. Liu, D. Chen, D. S. Weung, F. Zhang, X. Wang, "Face Liveliness Detection Using a Flash Against 2D Spoofing Attack", In Proceedings of IEEE Transactions on Information Forensics and Security, vol. 13, no. 2, pp. 521-534, 2018.
[11] M. Singh, A.S. Arora, "A robust anti-spoofing technique for face liveliness detection with morphological operations", In Proceedings of International Journal for Light and Electron Optics, vol. 139, pp.347-354, 2017.
[12] Patrick P. K. Chan, Weiwen Liu, Chien-Chang Hsu, "Face liveness detection using a flash against 2D spoofing attack", In Proceedings of IEEE Transactions on Information Forensics and Security, vol. 13, no. 2, pp. 521-534, 2018.
[13] M. Smiatacz, "Texture Features for the Detection of Playback Attacks: Towards a Robust Solution", In Proceedings of Journal of Advances in Intelligent Systems and Computing, vol. 977, pp. 214-233, 2020.
[14] Z. Boulkenafet, J. Komulainen, A. Hadid. "Face Spoofing Detection Using Colour Texture Analysis", In Proceedings of IEEE Transactions on Information Forensics and Security, vol. 11, no. 8, pp. 1818-1830, 2016.
[15] S. Tirunagari, N. Poh, D. Windridge, A. Iorliam, N. Suki, A. T. S. Ho, "Detection of Face Spoofing Using Visual Dynamics", In Proceedings of IEEE Transactions on Information Forensics and Security, vol. 10, no. 4, pp.762- 777, 2015.
[16] S. R. Arashloo, J. Kittler, W. Christmas, "Face Spoofing Detection Based on Multiple Descriptor Fusion Using Multiscale Dynamic Binarized Statistical Image Features", In Proceedings of IEEE Transactions on Information Forensics and Security, vol.10, no.11, pp. 2396-2407, 2015
[17] A. Pinto, H. Pedrini, W. R. Schwartz, A. Rocha. "Face Spoofing Detection through Visual Codebooks of Spectral Temporal Cubes", In Proceedings of IEEE Transactions on image processing, vol. 24, no. 12, pp. 4726-4740, 2015.
[18] L. Feng, L. Po, Y. Li, F. Yuan, "Face liveness detection using shearlet-based feature descriptors", In Proceedings of Journal of Electronic Imaging, vol. 25, no. 4, pp. 1-10, 2016
[19] J. Galbally, S. Marcel, J. Fierrez, "Quality Assessment for Fake Biometric Detection: Application to Iris, Fingerprint, and Face Recognition", In Proceedings of IEEE Transactions on Image Processing, 2014, vol. 23, no. 2. pp. 710-724, 2014.
[20] D.Wang, X.Tan, "Bayesian Neighbourhood Component Analysis", In Proceedings of IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 7, pp. 3140-3151, 2018.
[21] R. Gross, V. Brajovic, "An image pre-processing algorithm for illumination invariant face recognition", In Proceedings of 4th International Conference on Audio-and Video-Based Biometric Personal Authentication, pp. 10-18, 2003.
[22] R. J. Raghavendra, R. Sanjeev Kunte, "Extended Local Ternary Pattern for Face Anti-spoofing", in Proceedings of Lecture Notes on Electrical Engineering, Spinger, pp. 221-229, 2020.
[23] R. J. Raghavendra, R. Sanjeev Kunte, "Extended Local Ternary Co-relation Pattern: A novel feature descriptor for face Anti-spoofing", in Proceedings of Journal of Information Security and Applications, vol. 52, pp. 1-10, 2020.
[24] Maatta, J., Hadid, A. and Pietik, M, " Face spoofing detection from single images using micro-texture analysis", In Proceedings of International Joint Conference on Biometrics, pp. 1–7, 2011.
[25] X. Tan, Y. Li, J. Liu, L. Jiang, "Face liveness detection from a single image with sparse low rank bilinear discriminative model", In Proceedings of European Conference on Computer Vision, pp. 504–517, 2010.